Rediscovering the Historical Methodology of the Earth Sciences by Analyzing Scientific Communication Styles

Jeff Dodick — Science Teaching Center, Hebrew University of Jerusalem, Givat Ram, Jerusalem, Israel 91904
email: jdodick@vms.huji.ac.il

Shlomo Argamon — Department of Computer Science, Illinois Institute of Technology, 10 W. 31st Street, Chicago, IL 60616, USA
email: argamon@iit.edu

Abstract

Despite the still-regnant concept of science proceeding by a monolithic "Scientific Method", philosophers and historians of science are increasingly recognizing that the scientific methodologies of the historical sciences (e.g., geology, paleontology) differ fundamentally from those of the experimental sciences (e.g., physics, chemistry). This new understanding promises to aid education, where currently students are usually limited to the dominant paradigm of the experimental sciences, with little chance to experience the unique retrospective logic of the historical sciences. A clear understanding of these methodological differences and how they are expressed in the practice of the earth sciences is thus essential to developing effective educational curricula that cover the diversity of scientific methods. This chapter reviews the question of historical scientific methodology (focusing on geology), as it has been addressed by historians, philosophers, science educators, and working scientists. We present results of a novel linguistic analysis of scientific texts, which shows that such posited methodological differences are indeed reflected in scientific language use. Characteristic features of historical scientists' language can be directly connected to aspects of historical scientific methodology, as explicated by philosophers and historians of science. This shows that the same methodological concerns are reflected in working scientists' conceptualizations of their discipline. These results give guidance to science educators, in the light of the recent emphasis on teaching language skills, as in "Writing Across the Curriculum", to focus on teaching and evaluating language and discourse skills within the methodological conceptual framework of the historical sciences.

Introduction

There is a widely held misconception that scientists make use of a single, universal scientific method in their work (Cartwright, 1999; Cleland, 2001, 2002; Hacking, 1983, 1999). According to many modern science textbooks at both the high school and even university level, this method involves conducting experiments under controlled conditions in a laboratory to test hypotheses. Indeed, this belief was so strong at one time, that as Hacking (1983, p. 149) notes the "experimental method used to be just another name for the scientific method". However, not all hypotheses can be tested in the laboratory. Historical hypotheses postulate particular causes for currently observable phenomena based on the (uncontrolled) traces that they leave behind. Such historical hypotheses are connected to the earth sciences, but also play a role in fields as disparate as evolutionary biology, astronomy and archeology.

As Cleland (2001; 2002) notes, scientists are well aware of the differences between experimental and historical sciences with regards to the difference in testing hypotheses. Indeed, scientists in experimental fields have often disparaged the claims of their colleagues in the historical sciences, contending that the support offered by the evidence in such fields is too weak to count as "good science" (Cleland, 2001, 2002; Gould, 1986).

Such criticism has had a profound effect on science education, especially at the pre-college level (Dodick and Orion, 2003). To a large degree, the earth sciences are not represented as a standard part of the core science curriculum across the United States. In part, this may be due to the thinking that since geology does not conform to the supposed universal scientific methodology, it does not represent the type of science that should be taught in the pre-college curriculum. One example of this prejudice at the pre-college level can be seen in the letter of protest distributed by the American Geological Institute assailing the January 25, 2002 *Draft California Science Framework for K-12 Public Schools.* Under this plan all students in California were mandated to take 2 years of laboratory science. Earth Science was acceptable only "if they have as a prerequisite (or provide basic knowledge) of physics, chemistry and biology". (See Dodick and Orion (2003) for the full text).

In the case of the university, it cannot be said that historical sciences such as geology are missing from its science curriculum. Nonetheless, Gould (1986, p. 74) still sees subtle prejudice acting against such fields as he notes:

> Harvard organizes its core curriculum and breaks conceptual ground by dividing sciences into two major styles experimental-predictive and historical, rather than traditionally by discipline. But guess which domain becomes 'Science A' and which 'Science B'?

The overall affect of such discrimination both at the high school, and university level is that students often lose the chance to gain scientific literacy not only in earth science, but in many historically based subjects which, by definition, require an understanding of temporally related changes.

However, things are changing on the education front. The recent national standards documents (American Association for the Advancement of Science, 1990, 1993; National Research Council, 1996) have called upon teachers to convey the diversity of scientific methods, including the historical method. To support this goal it will be necessary to create new learning materials that more accurately represent the different methodologies of science. More importantly, it will be necessary to change the philosophical mind set of teachers so that they understand how science can accurately investigate the past using methodologies provided by the historical sciences.

As part of this effort, there is growing interest in examining the nature of the methodological differences between the experimental and historical sciences. Recently, philosophers of science as well as science educators have started to examine such differences (Cleland, 2001; 2002; Cooper, 2002; 2004; Diamond, 1999; Frodeman, 1995; Gould, 1986; Raab and Frodeman, 2002; Rudolph and Stewart, 1998). Such work is an important first step in that it establishes the fact that indeed there is no single logically superior method in science, and hence historical methods are not inherently inferior to those used in experimental science. More importantly, efforts are underway to elucidate more precisely what the historical methods are, and how such methods logically follow from the kinds of questions upon which historical scientists work.

In this chapter, we review the question of methodology in the historical sciences (with a focus on geology) as it has been addressed by philosophers, science educators, and working scientists. We then present a logical next step, which is the empirical investigation of these methodological questions. This is addressed here via the analysis of scientists' language, in the light of the recent emphasis on improving language skills in the disciplines, as in the

"Writing across the Curriculum" program (Emig, 1988; Locke, 1992; Klein and Aller, 1999; Moore, 1993)

Our contention, borne out by our research results, is that methodological differences between scientific fields are reflected in related differences in language use in scientific communication. Educators should therefore be aware of the features characterizing such language use, both to be able to draw students' attention to those features, as well as for evaluating student writing about the sciences. By thus elucidating linguistic features that are specifically associated with the earth sciences, we hope to point the way towards new language-oriented foci for earth science education.

We have begun this work by analyzing stylistic variation in language use between scientists in different fields. Our initial results comparing writing style in peer-reviewed articles from Paleontology, Evolutionary Biology, Geology, Physics, Physical and Organic Chemistry give promising results which show how stylistic analysis of the ways in which different kinds of scientists organize their communication can provide empirical evidence supporting the existence and importance of the methodological distinctions between fields posited by philosophers and historians of science.

Perspectives from the history and philosophy of science

As noted, some scientists (mistakenly) disparage the method of historical science as being less scientific than the experimental method. In part, this attitude can be traced to developments in the philosophy of science during the 19th and early 20th centuries (Rudolph and Stewart, 1998). Based on the success of Newtonian physics, philosophers took physics as the model of how proper science should be done. Indeed, the views of the 19th century philosophers on science reflected a strong bias towards direct observations made during (controlled) experimental manipulation of nature (Mayr, 1985; Kitcher, 1993; Rudolph and Stewart, 1998). It was mistakenly believed at the time that by using experimental methods, that is, by manipulating independent variables and measuring changes in dependent variables, scientific conclusions could be established with greater certainty.

Citing the case of Charles Darwin, Rudolph and Stewart (1998) and Gould (1986) note that many in the British scientific community who accepted Darwin's conclusions concerning evolution by Natural Selection still had deep misgivings about the historical-comparative method he used to arrive at such conclusions. In part, as Rudolph and Stewart (1998) note, such misgivings were based on the philosophical backdrop to this period. Earlier in the 19th century, philosophers including John Herschel in his ***Preliminary Discourses on the Study of Natural Philosophy*** (1830) and John ***Stuart Mill's System of Logic*** (1843) were establishing the supposed foundations of scientific methodology (Hull, 1973). According to them it was firmly associated with the empirical, inductive methods of Newton (and other physicists).

This restrictive account of science, however, ignored the scientific basis of evolutionary biology up to the time of Darwin. Naturalists of the time described and classified the phenomena of the past and present, which might explain general patterns of development. However, it was not until Darwin that someone was able to provide an evolutionary mechanism that was accepted by the scientific public. In contrast to the complex diversity of life that naturalists tried to explain, with their systemic methodology, the methods of Newton attempted to reduce the world to its simplest forces. The success that Newton had with such reductionist methods provided the model for science that many in the philosophical community of the 19th century used as their standard for producing *certain* knowledge.

This bias of philosophers towards Newtonian physics as the model of science continued well into the 20th century, and was reinforced by the development of the school of thought known as 'logical empiricism'. Based upon this view, the goal of inquiry in science was the establishment of laws or universal mathematical statements (Sober, 1993). In turn, such sciences produce theories that lend themselves well to experimental confirmation

(Rudolph and Stewart, 1998). While such a methodology can be, and is, generally applied in the physical sciences, attempts at re-configuring historical sciences such as evolutionary biology and geology towards this methodology have largely failed (as we shall see).

As a result of this bias against historical sciences in general, some scientists and philosophers of science have mistakenly labelled geology as a "derivative science" (see Schumm (1991) as an example of this "derivative" claim). In this view, geology is seen as a synthetic science whose reasoning could literally be reduced to elements supplied by physics and chemistry. Moreover, in contrast to experimental sciences in which all variables could be controlled, geology was said to have a series of problems (including its great expanse of geological time, and gaps in the stratigraphic record) that undercut its claim to knowledge. Indeed, this situation is reflected by the meagre attention that the earth sciences have received by historians and philosophers of science, themselves. This neglect was based largely on the assumption that since geology is a synthetic science, which is easily subsumed under the (experimentally based) physical sciences, there is little need to study its historical development.

Paralleling these developments in philosophy, we also find that the geological community has itself sometimes contributed to the understanding that geology (as a historical science) is subservient to (the experimental science of) physics. To better understand this claim we will briefly review the history of the "founders" of (British) geology, James Hutton and Charles Lyell. (This history has been detailed elsewhere, including Dodick and Orion (2003) and so by necessity this review will be brief). The salient point here is that this history has greatly influenced the way modern geology (as an historical science) is both practiced and perceived both within and without the field.

James Hutton (1726-1797) is sometimes portrayed as the "father of geology", largely due to his logical deduction of the nature of geological time. Before Hutton, it was understood that the earth landmass would ultimately wear away due to erosional processes. However, this understanding clashed with Hutton's belief in a purposeful world created by God, which might sustain favorable living conditions for man (Toulmin and Goodfield, 1965; Albritton, 1980; Gould, 1987; Burchfield, 1998). Based on this belief, Hutton viewed the earth as a self-maintaining machine, continuously cycling through processes of deposition, and uplift so as to counter the effects of erosion and thus maintain environmental stability. With such continuous, cycles, geological time was deemed by Hutton to be an endless, immeasurable entity.

Hutton's depiction of geological time did represent an advance on the biblically determined age of 6000 years. However, he ignored the unique element of history within his model of geological time, in that he believed that the forces shaping the Earth's surface in the past were the same as those which shaped it in the present (Rudwick 1985; Gould, 1977; Goldman, 1982; Gould, 1987; Laudan, 1987). With such perfect, repeating cycles operating, there was no room for true change or progression.

Not surprisingly, given the philosophical tenor of the time, a major factor in Hutton's denial of historical progression was Sir Issac Newton's model of a mechanical universe (Laudan, 1987). Such a universe was ruled by a system of laws in which the planets eternally circled the Sun in timeless perfection. This influence was so great that Hutton copied Newton's theoretical language when writing about the balanced set of forces which drove his geological cycles (Laudan, 1987).

Hutton's theory attracted relatively few followers amongst fellow geologists because they rejected both his interpretation of forces, as well as the "entirely causal" nature of his model (Laudan, 1987, p. 134), which argued against the historical nature of earth's development. Indeed, Hutton showed little interest in historical geology; this is corroborated by the fact that he gave almost no attention to the fossils, which were becoming important in unravelling the progression of earth history (Gould, 1987).

Nonetheless, Hutton's theory became widely disseminated through John Playfair's (1802) publication, ***Illustrations of the Huttonian Theory of the Earth***. As a physicist, Playfair treated Hutton's theory as a branch of Newtonian physics, while eliminating its theological implications (Laudan 1987; Dean, 1992; Rudwick, 1998). As we have noted previously, Newtonian principles cast a long shadow over many Enlightenment ideas about the natural world, so Playfair's adoption of them specifically within geology was natural. Playfair's treatment of Hutton was especially attractive to Charles Lyell (1797-1875), who incorporated Hutton's ideas in one of the most important works of 19th century geology, ***The Principles of Geology*** (Lyell, 1830-1833).

In The Principles, Lyell provided a methodological outlook for geological questions (labelled the doctrine of uniformitarianism by philosopher William Whewell (1794-1866)) which demanded that geologists assume (apriori) that actual causes observed in the present were wholly adequate to explain the geological past, not only in kind but also degree (Rudwick, 1998). Based on this view, Lyell saw the earth as a dynamically balanced, steady state system, in which change was gradual and continuous, but led nowhere (Rudwick 1970; Rudwick, 1985; Gould, 1987; Burchfield, 1998). Thus, Lyell had a vague notion that geological time was vast, but "his notion of the earth's dynamics was curiously atemporal" (Burchfield, 1998, p. 139). In contrast, most geologists favoured a progressive earth history based on the overwhelming evidence of fossiliferous strata that were systematically being classified during the 19th century (Bartholomew, 1976, Toulmin and Goodfield, 1965).

Lyell required such a restrictive methodological outlook because he believed that the way to avoid inconsistency in geology was through a strict adherence to logic. Specifically, this meant the adoption of the two scientific principles of Vera Causa and enumerative induction (Laudan, 1987; Baker, 1998; Rudwick, 1998). The source of the former, not surprisingly was Newton (and his mathematical treatise ***Principia***) in which he argued that in science one must refer only to those existing causes which were sufficient to produce an effect. In the case of the latter, enumerative induction is a pattern of scientific reasoning in which the collection of facts takes precedence, unsullied by any theoretical presuppositions, from which theory might then emerge (Hull, 1973; Laudan, 1977). G.B. Greenough, the first President of the Geological Society of London, urged his members to adopt this strict form of induction, as it was "unassailably scientific" (Laudan 1987, p. 168).

Lyell's adoption of these two principles was his response to geologists who referred to catastrophic events as forces that shaped the earth's past. Such catastrophism was an anathema to Lyell because it implied that geology relied upon unknown causes, thus violating the logical principle of simplicity (which states that the best scientific explanations are those that consist of the fewest assumptions). Lyell believed that the apriori application of uniformity (based upon Vera Causa) was necessary, if geology, like physics was to be considered a valid, logical science (Baker 1998; 2000).

William Whewell attacked Lyell's adoption of such principles as unsuitable for a science such as geology. Whewell (1837) believed that geology was concerned with "the study of a past condition, from which the present is derived by causes acting in time". Thus it was inappropriate to apply Vera Causa, towards specifying the nature of those causes, a priori. Moreover writing about Lyell's inductive logic, Whewell (1837, v.3. p. 617) notes:

> (Lyell's) 'earnest and patient endeavour to reconcile the former indication of change', with any restrictive class of causes-a habit which he (Lyell) enjoins-is not the temper to which science ought to be pursued. The effects must themselves teach us the nature and intensity of the causes which have operated.

Simply put, Whewell was denying that the kind of induction advocated by Lyell, which he had borrowed from physics, was appropriate to a science such as geology (Baker 1998).

Nonetheless, Lyell's uniformitarian view has had a deep effect on geology in the modern age (Ager 1981; Hsu 1994; Baker 1998). For this chapter, we will briefly document two controversies that appear to clash with this notion of Lyellian uniformitarianism:

1. Based on uniformitarian logic it has been proposed in the past that during the last ice age many landforms in North America and Eurasia were formed exclusively by the slow, gradual action of glaciers. Thus in 1923 when Bretz (1923) proposed that a catastrophic flood was responsible for the origin of the Channeled Scabland region of eastern Washington (state) it ignited a vehement controversy in the geological community that extended over decades. Indeed, it was not until the late 1960s that enough evidence had accumulated for this hypothesis to became acceptable to many geologists; the discovery of features, such as giant ripple marks, as well the identification of the source of the floods themselves-the Late Wisconsin glacial Lake Missoula-caused many to rethink their opposition to a catastrophic flood as the ultimate origin of the Scablands (Bretz, 1969; Baker, 1978; Pardee, 1940; 1942). A subsequent chapter to this debate was added by Wait in 1985 when he established that glacial Lake Missoula drained periodically as scores of colossal jokulhlaups (glacier outburst floods), based on his discovery of more than 40 successive flood laid rhythmites which had accumulated in back-flooded valleys in Southern Washington. More recently, in 1996, researchers working in Iceland actually witnessed such jokulhlaups in action, while documenting their effects on the surrounding landscape (Bjornsson, 1998; Roberts, et. al., 2001). Certainly such evidence should convince those who adhere to Lyell's interpretation of uniformitarianism that such catastrophic flooding was both a strong and persistent influence on ice age landforms.

2. Although more controversial, the debate concerning the mass extinction at the end of the Mesozoic has also exposed (at least according to some earth scientists) the flaws with (Lyellian) uniformitarian thinking (Hsu, 1994; Gould, 1986; Glen, 1994). For some earth scientists, the idea that such extinction was due to the impact of a massive extraterrestrial body was impossible because it violated their apriori prejudice against large-scale catastrophic events, which have not been documented in the present day. However there is increasing geological evidence supporting this theory. The Chicxulub crater discovered in Yucatan, Mexico in 1981 appears to be contemporaneous with the Mesozoic extinctions, and is evidence of an impact that appears to be forceful enough to have caused a major environmental calamity (Hildebrand, Kring and Boyton, 1991). Moreover, such impacts are not as rare as once thought. Approximately, 150 craters have now been documented on earth, most of which are less than 200 million years old (Grieve, et. al., 1995), with at least some of these impacts arriving as multiple events (e.g. Spray, Kelley, and Rowley, 1998). Most importantly, as with our example of catastrophic flooding, evidence for the effects of such impacts have been now been witnessed in real time, as seen in the dramatic collision of disrupted comet Shoemaker-Levy 9 with Jupiter in 1994 (Orton, et. al., 1995).

These two examples emphasize that a key part of understanding causal process in geology is found by interpreting the landforms, structures and rocks which make up the earth (Baker 1998; 2000; Turner, 2000). In simple terms, the student of geology will find his answers in the concrete historical materials of the field. As Ager (1981), Baker (1998) and Hsu (1994) argue, Lyell's interpretation of uniformitarianism with its emphasis on creating a logically valid science of geology may be a historical relic with little place in modern earth sciences. Indeed, regarding the pervasive influence of Lyell's uniformitarianism, Hsu (1994,

p. 218) notes that a study he himself conducted on ocean chemistry, was met with skepticism by many of his colleagues because as they said it "contradicts Lyell's uniformitarianism, the fundamental principle of geology".

Current state of affairs in science education

The neglect of the earth sciences (and other historical sciences) is also reflected within the modern school curriculum. To understand how this developed, we will briefly focus on the history of the earth sciences in the (American) school system.

Prior to the twentieth century, geology was an intrinsic part of the high school curriculum for college bound students in the United States. In part, this reflected the fact that geologists held pre-eminent roles in American science, and wielded their political power accordingly. However, by 1910 such courses had low enrolment, and geology had effectively become an elective (Orion et al., 1999). The conventional wisdom of the time regarded physics and chemistry as having greater social importance than geology because they aided students in developing problem solving skills (Bybee and DeBoer, 1994).

In the early 1960's, the domination of the physical sciences within the American education system was bolstered by the supposed national security threat posed by the Soviet space program. Science education focused on "the logical structure of the sciences and the processes of the sciences" (Bybee and DeBoer 1994, p. 373) with less emphasis on its personal and social applications. Thus, the sciences which were emphasized, physics and chemistry, were those needed to maintain the United States' (military and) technological advantage over the Soviet Union (Mayer and Kumano, 2002).

Towards the very end of the 1960's, a program in the earth sciences reappeared in the American school system, the ESCP (Earth Science Curriculm Project) (American Geological Institute 1967), with a second program following in the early 1970s, CEEP (Crustal Evolution Education Project). In theory, the implementation of such curriculum projects suggested that the earth sciences had achieved equal status with biology, chemistry and physics. However, this was not the case; although the ESCP received positive evaluations (Champlin, 1970), it does not seem to have made a lasting impact in the American secondary school system. CEEP was specifically designed to incorporate the new paradigm of plate tectonics, but was never successfully implemented in part due to "a lack of interest in the topic" (Orion et al. 1999).

Concurrently, these units gave the somewhat distorted image that the focus of geology was on the physical rather than the historical. This trend has continued until today. Both the Activity Sourcebook for Earth Science (Mayer et al., 1980) and later Earth Science Investigations (Oosterman and Schmidt, 1990) contain material found in the ESCP and CEEP, with the majority of the activities being strongly weighted towards the physical side of the earth sciences. This is all the more puzzling when one considers the fact that:

> These collections are the most widely accepted examples of what scientists and science teachers through their professional organizations have felt important to use in teaching earth science problem solving (Ault, 1994, p. 270).

In other words, even amongst many professional earth scientists and earth science educators, there is a bias towards a geology that is strongly physical, rather than historical. However, if this is truly the case, one might question why it is necessary to teach the earth sciences, if it simply replicates many physics based experiences?

In the last decade, a major movement for change in science education began; this movement is not only affecting the earth sciences, but the entire pre-college science curriculum. One of the key changes called for in these reforms concerns the style of learning done in the classroom, with a major focus on learning materials that emphasize inquiry. (Pre-college this includes the American Association for the Advancement of Science's well known Project 2061; at the college level this includes NSF grant 96-139, Shaping the Future of Undergraduate Education). There are many definitions of this term, but for our purpose, inquiry based materials are those that more closely reflect the type of investigations that are pursued in authentic science environments. The understanding is that by replicating such environments, students will not only learn more content, but will also develop the necessary critical thinking skills to pursue independent work in science.

This call for new education materials is partly a reaction to the typical experience of pre-college students who often experience rote memorization in their science lessons. Moreover, if they experience any other learning environment, aside from lecture, it is usually a cook-book lab experience in which the results are pre-determined. We might add that such practices have often emphasized the prevailing myth amongst both students and teachers that the experimental method is the single universal method of science, because it is rare for students to venture beyond their classroom into the field where most historical sciences are pursued.

However, the recent reforms proposed in the national standards documents (American Association for the Advancement of Science, 1990, 1993; National Research Council, 1996) call upon teachers to convey the diversity of scientific methods, including the historical method. This is a positive step in that at least some of the student population touched by such reforms might be more attracted to historical sciences, rather than their experimental counterparts. This has certainly been the case with some students in Israel, where a program in the geosciences (at the high school level) has been offered nationally since the late 1990's. If one of the goals of science education is to increase scientific literacy amongst future citizens, then offering such choice today is an excellent policy move.

Further, many historical sciences, such as the earth sciences have an advantage in that they integrate the conceptual skills of the other major sciences; this reflects a general trend in the sciences toward more interdisciplinary cooperation between different fields and practitioners. So again, if the goal is for students to experience authentic science environments, then being exposed to an interdisciplinary field such as the earth sciences has strong pedagogic value.

Finally, it might be noted that another goal of the current reform movement in science education is to create materials that are relevant to the lives of today's students. One important topic that fits this goal is the deepening environmental crisis. As both Turner (2000) and Frodeman (1995) have argued the systemic and historical methodology of the earth sciences has much to contribute towards educating students about how human activity has impacted upon the natural environment, and how to solve such environmental problems.

Currently, the design of new curriculum materials to meet the needs of the reforms is in its first stage of development. Thus, it is still too early to evaluate what effect such reforms will have on both the perception and teaching of the earth sciences within the pre-college science curriculum (Orion et al., 1999) If present experience is any indicator there is much work to do to overcome the ingrained prejudices that educators have against historical science in general and earth science specifically. For example, even today the guidelines for many science fair projects (a mandatory school requirement in many of the States) require the use of the experimental method, and thus naturally exclude many projects of a historical nature, including the earth sciences.

Geology as a historical science: A reconsideration

Recently, a growing group of philosophers, scientists, and educators have become sceptical about the existence of a single method for all of science (Cleland, 2001, 2002; Cartwright, 1999; Cooper, 2002, 2004; Diamond, 1999; Gould 1986; Hacking, 1983, 1999; Mayr, 1985; Rudolph and Stewart, 1998). It is now recognized that experimental methods and evidential reasoning can play different roles in different sciences (Hacking, 1983).

Nonetheless, as Cleland (2001; 2002) notes (most) philosophers have been reluctant to make large scale comparisons concerning methodology across different disciplines. The ultimate result is that many outside the philosophy of science are often left with the mistaken impression that historical science is lacking when compared to experimental science.

While details of approaches and focus differ among these different researchers, a broad outline of the difference between experimental and historical sciences may be given. Our formulation here considers four interrelated fundamental dimensions categorizing this difference (modified from Diamond (1999, pp. 421-424)):

- Is evidence gathered by manipulation or by observation?
- Is research quality measured by effective prediction or explanation?
- Is the goal of the research to find general laws or statements or ultimate (and contingent) causes?
- Are the objects of study uniform entities (which are interchangeable) or are they complex entities (which are each unique)?

Experimental science, as is well known, gathers knowledge by controlled experimentation, in which natural phenomena are manipulated in order to test a theory. The quality of such a theory is measured by consistency of its predictions with such experiments, and ideally, such a theory expresses a general causal law or universal statement that is applicable to a wide variety of phenomena in many contexts. Finally, the form of such research is dictated in large part by the study of uniform entities, either singly or in assemblage; the fact that such entities (atoms, molecules, genes) are identical, or nearly so, makes the formulation of general laws possible in principle, and experimental reproducibility a reasonable requirement in practice.

Historical science, on the other hand, investigates ultimate causes which often lie very deep in the past, and whose effects are observed only after very long and complex causal chains of intervening events. Consequently, evidence is gathered by observation of naturally occurring traces of phenomena, since manipulation is impossible (e.g., we cannot wait millions of years for the results of a hypothetical geological experiment!). Comparison of similar phenomena has thus developed as a key analytical technique, as typified by the kind of 'natural experiment' (described by Diamond (1999)) where systems differing in a key variable are compared in order to estimate the effect of that variable on other system characteristics.

This focus on investigating past causation further implies that the ultimate test of quality in historical science is explanatory adequacy, because the phenomena under investigation are unique and contingent, and so cannot be usually expected to repeat in the historical record. The methodology of such explanatory reasoning derives from what Cleland (2002) calls the 'assymmetry of causation', in that effects of a unique event in the past tend to diffuse over time, with many effects being lost and others muddled by other intervening factors. Making sense of this complexity requires, therefore, 'synthetic thinking' a lá Baker (1996), in which one fits together complex combinations of many pieces of evidence to form arguments for and against competing hypotheses.

In addition to sorting through this great causal complexity, historical scientists must also deal with the complexity of the individuals under study. Unlike (subatomic) particles (for example), which are all uniform, the entities studied by historical scientists—people, species, strata—are all unique (though often similar) individuals, whose precise details of configuration and function are not always recoverable, in practice or even in theory. This eliminates the possibility of formulating universal laws of behavior, allowing only statistical statements of relative likelihoods at best—it is very difficult to rule a specific possibility out entirely, but rather arguments for and against hypotheses must be made on the preponderance of the best evidence. Thus reasoning about the relative likelihood of different assertions is endemic to historical science's synthetic thinking.

Studying Scientific Communication Styles

As we have seen, despite the historic downgrading of historical methodology in science, strong arguments can now be made that the historical sciences have, and indeed require a different type of methodology from that used in experimental sciences such as physics. Clearly understanding these methodological differences will be key to effectively increasing the role historical sciences such as geology play in secondary school and university science curricula. However, gaining such an understanding requires empirical investigation of how scientists actually do science in practice, with one major element of doing science being how scientists communicate; the ways in which scientists reason, make discoveries, and persuade their colleagues are all reflected in how they communicate with one another.

This empirical approach differs radically from more 'traditional' methods of studying scientific methodology based on either scientists' own introspection (such as Karp, 1989) or philosophical-historical analysis of scientific discovery (for example Kuhn, 1962, and Popper, 1959). Here the aim is to study what scientists say or write while doing science, in order to understand what sorts of cognitive and reasoning structures they actually use in practice, not in rationally reconstructing what they did.

The study of communication as a way to investigate scientific cognition has been most recently addressed in the experimental sciences, from the perspective of cognitive science by, amongst others, Dunbar and colleagues, in studies of research activities in molecular biology laboratories (Dunbar, 1995; Dunbar, 1999; Dunbar and Blanchette, 2001), and from an anthropological-linguistic perspective by Ochs and Jacoby (1997) and Ochs, Jacoby and Gonzales (1994), in their studies of physicists' talk in the classroom. However, little work has been done on how historical scientists communicate, and no research, to our knowledge, has addressed possible differences between communication modes in historical and experimental sciences. We contend that such empirically-based studies are essential to properly understanding the characteristic communication and reasoning modes of the historical sciences (in comparison with the same in experimental sciences).

There are a number of different types of communication between scientists (in any field), which can be roughly classified by the mode of communication (speech or formal/informal writing) and by audience for that communication (research collaborators, other colleagues, and students), as laid out in Table 1.

Insert Table 1 about here

The study of speaking patterns in scientists will require videotaping and transcription, as a prelude to linguistic discourse analysis. This is expensive and time-consuming, though, and generally precludes working with very large data sets. Written material, when available in electronic form, is often easier to work with for analysis, however, since automated tools can more easily be brought to bear. This is particularly true for formal written texts, such as journal articles or textbooks. While analysis of such formal written texts cannot give us direct insight into how scientists reason while doing historical based field work, or developing arguments for or against a theory, they do give a unique window into how scientists *present* their ideas, i.e., what underpinning conceptual structures they see as essential to communicating their ideas with their peers.

Our key idea here is that many aspects of language style directly reflect how people organize ideas, argue for or against a position, or explain new concepts. Specifically, in scientific communication, we hypothesize that different scientific methodologies should imply concomitant differences in how scientists organize their discourse, and that these differences in turn will be reflected in their use of language. Thus, while textual analysis can say little to any epistemological concerns, differential analysis of scientific language use may provide us with empirical evidence for and against different proposals for describing scientific methodology in different fields. Furthermore, if such evidence is positive, showing meaningful linguistic differences between fields with different posited methodologies, the linguistic characteristics specific to each particular field may then provide useful guidance to educators in teaching and evaluating language and reasoning skills in that field.

To make this concrete, for the case of experimental and historical sciences, consider again the dimensions of methodological difference discussed above in the previous section (Geology as a Historical Science) which together generate the following basic hypotheses regarding variation in language use between scientists in historical and experimental sciences:

- First, we expect historical science's posited focus on observation and explanation to be expressed by language containing more explicitly comparative language than that of experimental science. On the other hand, since experimental science is posited to rely more on manipulation of variables in nature and is focused on consistency with prediction, we would expect its language to more often qualify assertions as to their predictive value or consistency with predictions.

- Second, the focus of historical science on explanation in order to find ultimate causes implies that its argumentation will involve combinations of multiple lines of evidence of differing validity and generality, due to the complexity of the phenomena and the assymetry of causation (Cleland 2002). This further implies that historical science will use more complex and explicit qualifications of likelihood. In experimental science, on the other hand, we expect to see this less, since arguments will tend to be more narrowly constructed around a specific causal phenomenon; however, we might expect to see in experimental science qualifications related to possibility (can it happen?) or necessity (must it happen?).

- Third, the fact that historical scientists study complex and (ultimately) unique entities and systems by observation and comparison implies that they would need to express many separate pieces of information linked together by contextual or comparative links. For example, think of a geologist describing a particular site—connections between adjacent statements are likely to be geographical ("…and adjacent to that is…") or temporal ("when that process completed, then…"). Conversely, since experimental science focuses on deep causal descriptions of essentially uniform entities, we expect a more 'unifocal' prose, where links between assertion are based on tight causal, conditional, or temporal links (e.g., "X causes Y whenever Z").

We are currently developing a methodology for performing linguistic analysis of scientific texts (currently written, but we intend to later consider spoken interaction as well) that combines state-of-the-art computational techniques with Systemic Functional Linguistics (Halliday, 1994), to find evidence for claims such as the above. Systemic functional linguistics (henceforth, *SFL*) treats grammar as a structured set of options for realizing meanings via language in context; hence it is a useful theory for relating language structures to the rhetorical needs of the writer and reader. In our case, we wish to find out in what ways historical and experimental scientists structure their discourses differently, and to relate such linguistic phenomena back to the several dimensions posited to describe methodological variation between fields.

During the current phase of this research project, we are analyzing language use from a collection of peer-reviewed journal articles from both historical science and experimental science. The detailed composition of the corpus is summarized in Table 2.

Insert Table 2 about here

Our main goal is to find stylistic features of language use which reliably distinguish articles in one type of science from articles in the other. Stylistic features are often subtle features of word use (including syntax) related to the "connotation" of a text, rather than its more overt specific topic. For example, the frequency of a word like "of" gives a partial cue as to how a text modifies nouns, and hence how the author describes things (detail on relevant stylistic variables in our study is given below).

To obtain empirical evidence for methodological variation between the two fields under study, our stylistic analysis must answer the following two questions:

Q1. Can experimental science articles be reliably distinguished from historical science articles purely on the basis of stylistic features of the text? If they can, then we have prima facie evidence for differences in language use that may reflect methodological differences.

Q2. If they can be distinguished, what stylistic features are consistently significant for such classification? What structural properties of the text are indicated by these differing styles, and how do those relate to possible differences of methodology between the two fields (as expressed in the four fundamental dimensions categorizing historical and experimental sciences, mentioned previously)?

We reiterate here that we are not assuming that such a relationship between methodology and language exists, but rather we are testing such a hypothesis empirically, by analyzing original scientific texts.

Q1. Distinguishability

For our current corpus, we can answer the question of distinguishability between experimental and historical sciences in the affirmative, based on a computational study we recently conducted (Argamon and Dodick, 2004; 2005a; 2005b) on the corpus described above. We used a computer implementation of machine learning techniques which enabled efficient processing of large numbers of documents to find combinations of linguistic features (from a large predefined list) which gave the most accurate discrimination between different document groups (articles from different journals). (An excellent introduction to the field of machine learning is Mitchell (1997)). For those familiar with statistical analysis techniques, the methods we use have some similarity to discriminant analysis-though machine learning techniques have properties that are more desirable for our application here.

In brief, we defined a set of 93 relevant stylistic features (based on Matthiessen's (1995) SFL grammar of English) which could be effectively extracted automatically from the article texts. Each article was then processed and represented by a vector of 93 numbers, each one the relative frequency of a feature in that article. Once such 'feature-vectors' were created for the articles, we used the popular machine learning technique of Support Vector Machines (Joachims 1998) to find models for distinguishing articles in each journal from each other journal, based on these features. In our use of the method, the system constructs a linear weighting of feature values that distinguishes between two classes of training data. Accuracy of such learning was measured by the standard technique of 10-fold cross-validation, in which models are trained and tested on disjoint sets of the data. If experimental science writing is distinguishable from that in the historical sciences, we expect classification effectiveness to be higher for pairs of journals in different kinds of sciences than for pairs in the same kind, which we expect to be higher than for pairs of journals in the same specific field.

Figure 1 plots classification effectiveness (measured by cross-validation accuracy) for four classes of journal pairs: two journals in the Same field, different fields in either Experimental or Historical sciences, respectively, and journals in two Different kinds of sciences (Historical and Experimental). As the plot clearly shows, the greatest classification effectiveness (on average) is obtained for journals from different kinds of science, with intermediate effectiveness for journals from different fields in the same kind of science. Journals covering the same scientific field (more or less) are the least distinctive, as expected.

Insert Figure 1 about here

Q2. Features, structure, and methodology

We now turn to the question of which features are most relevant to stylistic differences between experimental and historical science writing. The stylistic features of a text which we consider here are drawn from three main categories based on Matthiessen's (1995) English grammar, which uses principles of Systemic Functional Linguistics (Halliday 1994). Each kind of feature we consider is connected to how scientists structure communication and relates to different propositions, as follows:

- **Cohesion** refers to how language can connect to its larger context (both within the text and outside of it, in the `real world'). The way in which cohesion resources are used in a text expresses how the author organizes ideas and relates them to each other. A key grammatical structure which contributes to textual cohesion is Conjunction, which describes different kinds of textual features which serve to link clauses together causally or logically. The first of the three main types of conjunction is Extension, which links clauses giving different information together, realized by words such as "and", "but", and "furthermore". The second type is Enhancement, which qualifies information in one clause by another (e.g., "similarly…" or "therefore…"). Third is Elaboration, which deepens a clause by clarification or exemplification (e.g., "in other words…" or "more precisely…").

- **Modality** qualifies events or assertions in the text according to their likelihood, typicality, or necessity. There are two main types: Modalization, which quantifies levels of likelihood or frequency (e.g., "probably", "might", "usually", "seldom"), and Modulation, which quantifies ability or necessity of performance (e.g., "ought to…", "should…", "allows…", "must…").

- **Comment** qualifies assertions in the text according to their rhetorical relation to the context. Eight types of comment are described by Matthiessen (1992): Admissive (clause is an admission: honestly...), Assertive (vigorous assertion: positively...), Presumptive (relates clause to presumptions: evidently...), Desiderative (relates to desirability: fortunately..., unluckily...), Tentative (clause is tentative: honestly...), Validative (scope of clause's validity: generally speaking..., in some cases...), Evaluative (evaluation of actors: foolishly..., justifiably...), Predictive (relates clause to predictions: as expected..., surprisingly...).

Detailed analyses of the features we found most indicative of experimental and historical science writing are given in Argamon and Dodick (2004; 2005a; 2005b); our main findings may be summarized as follows.

Within CONJUNCTION, we find *Enhancement* to be a consistent indicator for experimental science articles, while *Extension* was a consistent indicator for historical science articles. This is consistent with our hypothesis above that experimental scientists will tend to hew more closely to enhancing a single storyline, while historical scientists are more likely to discuss a greater variety of separate information elements (since observational methodology requires more detailed consideration of context).

Next, within MODALITY, we find *Modulation* to be a consistent indicator for experimental science articles, while *Modalization* was a consistent indicator for experimental science articles. This is consistent with our hypothesis above that experimental scientists will focus on the possible and the necessary, while historical scientists will more often explicitly qualify assertions by likelihood and frequency, due to the complex combination of multiple uncertain lines of evidence.

Finally, within COMMENT, we find *Predictive* comments to be consistent indicators for experimental science articles, while *Validative* comments were consistent indicators for historical science articles. This is consistent with our hypothesis above that experimental scientists will tend to structure their arguments around consistency of observations with theoretical predictions, while historical scientists will be more concerned with properly defining the scope of validity of various pieces of evidence and claims.

We thus see how examination of variation in features of language use between articles in different scientific fields may be used to support hypotheses about how scientists in these fields present their research and argue for (or against) scientific positions. Clearly, historical scientists are not merely attempting to copy experimental scientists, but rather are pursuing a different agenda, and for good reason. Beyond this basic point, though, these results suggest specific areas for educators to focus on when teaching students to read and write in the earth sciences. Explicit attention should be paid to how writers link descriptions of disparate elements of a landscape (say) into a synchronic or diachronic panorama, by use of conjunctions (both explicit and implied). The distinction between *Extensive* and *Enhancing* conjunctions can be brought out, to help students see the different functions of different kinds of clausal linkages. Furthermore, educators can show the need for caution in evaluating the validity of evidence, by drawing attention to writers' use of MODALITY and COMMENT.

Conclusions

Geology as well as other historical sciences, has in the course of history, sometimes been treated as a derivative science by scientists, philosophers, and educators. We contend that this image is misguided because it does not take into account the unique defining characteristics of such disciplines. In contradistinction to the experimental sciences, which tend to be predictive, and reductionist in orientation, historical sciences such as geology are descriptive, and systems oriented. To a large degree, such defining characteristics are dictated by the types of phenomena that historical scientists, such as geologists study, as well as the kind of reasoning that must be brought to bear to understand such phenomena. As we have shown, such reasoning is imbedded and reflected in the end products of earth science research-scientific communication (as seen in professional journal articles).

Due to the large number of variables associated with geological phenomena, geologists rely on an interpretive and narrative form of logic which is used to reconstruct such phenomena (Frodeman 1995; Raab and Frodeman, 2002; Turner, 2000). Such retrospective thinking often requires the geologist to make a meticulous survey of present conditions, which might then be used to explain phenomena of the past. Such historical and comparative explanations emphasize the differences between historical sciences, such as geology and experimental sciences, such as (experimental) physics, which primarily focuses on establishing time invariant laws or universal statements.

The discovery of geological time is a perfect example of this working methodology, in that it required the patient unravelling of uncounted numbers of fossil bearing strata in the field, before science recognized “deep time’s” full implication. Through such efforts, the earth sciences have provided a true sense of the earth’s historical development through the creation of a heuristic geological time scale. A single quantifiable number derived in the laboratory and describing the magnitude of the earth's age, although important to this story, could never have established this historical narrative of the earth's development. The revolution of "deep time" emphasizes geology's key role in reconstructing an image of the past. The environmental crisis, with its large collection of interconnected variables, emphasizes that the holistic, systemic and historical methodology of the earth sciences has much to contribute in future to both science specifically and the welfare of the planet in general.

In downgrading the historical methodology of the earth sciences, scientists have constructed an unnecessarily limited definition of science. Moreover, this definition has also meant that within education, students are necessarily limited to the dominant paradigm of the experimental sciences, with little chance to experience the unique retrospective logic of the historical sciences. This ultimately affects their understanding of subject areas outside of geology, including evolution, ecology and astronomy, which by definition require an understanding of temporally related changes. It also limits students' chances of being scientifically literate (a stated goal of current educational reforms).

To help rectify this situation, we follow in the footsteps of those scientists and philosophers and historians of science, including Baker, Cleland, Diamond, Frodeman and Gould, who elucidate specific hypothesized methodological differences between historical and experimental sciences, tracing these differences to the evidentiary requirements of the questions different scientists ask and the kinds of data they work with. We have taken their analysis a bit further to examine, empirically the conceptual underpinning of scientists' conceptions of their science. As our analysis of journal articles has shown, comparative linguistic analysis of peer-reviewed journal articles can lend empirical weight to such philosophical arguments, by showing how posited methods and conceptual structures are reflected in scientific language. We thus can refine and give more precise characterizations of the kinds of methodological reasoning used in different scientific fields.

Our use of automated machine learning techniques in computational stylistics enables analysis of potentially large collections of scientific texts. The continuation of this work will involve analysis of larger collections of texts from a wider variety of fields. Educational relevance will be gained by 'longitudinal studies' of student-written texts at different educational levels (secondary, undergraduate, graduate) as well as textbooks and journals. Such work has therefore the potential to influence the education process, by providing curriculum designers, and education researchers the tools to better design new learning materials that are consonant with the calls for learning experiences that reflect the methodological diversity of all of the sciences. In this way, students not only learn the skills of specific sciences, but also the "nature" of the specific sciences (a much desired component of modern curricula), and thus banish the misconception that the historical sciences are derivative or inferior to the experimental sciences. The ultimate result is that greater options are offered to students who enjoy the field based methods of historical sciences, such as geology. We believe that our research program for developing precise characterizations of how scientific language use develops over time, and differs between fields, will provide a concrete basis for elucidating conceptual issues in scientific methodology.

Moreover, such research could help create a new generation of assessment tools that will permit better evaluation of students at all levels of the education process, by focusing on how and when they develop the necessary writing skills to participate in inquiry based science, that replicates a more authentic research based learning experience. Indeed, the development of alternative forms of assessment that goes beyond simple "pencil and paper" tests which often examine memorization skills, is a major component of the current educational reforms (American Association for the Advancement of Science, 1993; National Research Council, 1996). Thus, the insights we gain into scientific communication will not only aid the cause of the historical sciences, but will in fact serve the purposes of all subjects in the science curriculum.

REFERENCES CITED

Ager, D.V., 1981, The nature of the stratigraphic record (2nd edition): London, Macmillan Press, 122 p.

Albritton, C.C., 1980, The Abyss of Time: Farrar and Strauss: New York, 256 p.

American Association for the Advancement of Science. 1990, Science for all Americans: New York, Oxford University Press, 418 p.

American Association for the Advancement of Science, 1993, Benchmarks for science literacy: New York, Oxford University Press, 272 p.

American Geological Institute, 1967, Investigating Earth Systems: Boston, Houghton-Mifflin Co.

Argamon, S., and Dodick, J.T., 2004, Linking rhetoric and methodology in formal scientific writing, *in* Proceedings, 26th Annual Meeting of the Cognitive Science Society, Chicago, IL, August 2004, p. http://www.cogsci.northwestern.edu/cogsci2004/sessions.html.

Argamon, S., and Dodick, J.T., 2005a, The languages of science: A corpus-based study of experimental and historical science articles, *in* Proceedings, 27th Annual Meeting of the Cognitive Science Society, Stresa, Italy, August 2005, p. http://www.cogsci.northwestern.edu/cogsci2004/sessions.html

Argamon, S and Dodick, J.T., 2005b, A corpus based study of scientific methodology: Comparing the historical and experimental sciences, *in* Shanahan, J.G., and Qu, Y., and Wiebe, J., eds., Computing attitude and affect in text, Dordrecht, Netherlands, Springer (in press).

Ault, C.R., 1994, Research on problem solving: Earth science, *in* Gabel, D.L., ed., Handbook of research on science teaching and learning: New York, Macmillan Publishing, p. 269-283.

Baker, V.R. 1978. The Spokane flood controversy and the Martian outflow channels: Science, v. 202, p. 1249-1256.

Baker, V.R., 1998, Catastrophism and uniformitarianism: Logical roots and current relevance in geology, *in* Blundell, D.J., and A.C., Scott, eds., Lyell: The Past is the Key to the Present: Geological Society of London Special Publications 143, p. 171-182.

Baker, V.R., 2000, Conversing with the earth: The geological approach to understanding, *in* Frodeman, R., ed., Earth matters: The earth sciences, philosophy and the claims of the community: Upper Saddle River, NJ, Prentice Hall, p. 2-10.

Bartholomew, M., 1976, The non-progress of non-progression: Two responses to Lyell's doctrine: British Journal for the History of Science, v. 9, p. 166-174.

Bjornsson H, 1998, Hydrological characteristics of the drainage system beneath a surging glacier: Nature, v. 395, p. 771–774.

Bretz, J.H. 1923. The Channelled Scablands of the Columbia plateau: Journal of Geology, v. 38, p. 617-649.

Bretz, J.H. 1969. the Lake Missoula floods and the Channelled Scablands: Journal of Geology, v. 77, p. 505-543.

Burchfield, J.D. 1998, The age of the earth and the invention of geological time, *in* Blundell, D.J., and A.C., Scott, eds., Lyell: The Past is the Key to the Present: Geological Society of London Special Publications 143, p. 137-143.

Bybee, R.W., and DeBoer, G.E., 1994, Research on goals for the science curriculum, *in* Gabel, D.L., ed., Handbook of research on science teaching and learning: New York, Macmillan Publishing, p. 357-387.

Carman, A.J., Zhang, L., Liswood, J.L., and Casey, S.M., 2003, Methylamine adsorption on and desorption from Si(100): Journal of Physical Chemistry B., v. 107, p. 5491-5502.

Cartwright, N, 1999, The dappled world: A study of the boundaries of science, Cambridge, Cambridge University Press, 247 p.

Cleland, C. 2001, Historical science, experimental science, and the scientific method: Geology, v. 29, p. 987-990.

Cleland, C. 2002, Methodological and epistemic differences between historical science and experimental science: Philosophy of Science, v. 69, p. 474-496.

Champlin, R.F., 1970, A review of the research related to the ESCP: Journal of Geological Education, v. 18, p. 31-39

Cooper, R.A., 2002, Scientific knowledge of the past is possible: Confronting myths about evolution and the nature of science: The American Biology Teacher, v. 64, p. 476-481.

Cooper, R.A., 2004, Teaching how scientists reconstruct history: Patterns and processes: The American Biology Teacher, v. 66, p. 101-108

Dean, D.R., 1992, James Hutton and the history of geology: Ithaca, NY, Cornell University Press, 303 p.

Diamond, J., 1999, Guns, germs and steel: The fates of human societies. New York, W.W. Norton, 480 p.

Dodick, J. T., and Orion, N., 2003, Geology as an historical science: Its perception within science and the education system: Science and Education, v. 12, p. 197-211.

Dunbar, K., 1995, How scientists really reason: Scientific reasoning in real-world laboratories *in* Sternberg, R.J., and Davidson, J., eds., Mechanisms of insight: Cambridge, MA, MIT Press, p. 365-395.

Dunbar, K., 1999, The scientist In-Vivo: How scientists think and reason in the laboratory, *in* Magnani, L., Nersessian, N., and Thagard, P., eds., Model-based reasoning in scientific discovery. New York, Plenum Press, 89-98.

Dunbar, K., and Blanchette, I., 2001, The in-vivo/in-vitro approach to cognition: The case of analogy: Trends in Cognitive Sciences, v. 5, p. 334-339.

Emig, J. 1988, Writing as a mode of learning, *in* Tate, G., and Corbett, E.P.J., eds., The writing teachers source book (2nd ed.), New York, Oxford University Press, p. 85-93.

Frodeman, R., 1995, Geological reasoning: Geology as an Interpretive and historical science: Geological Society of America Bulletin, v. 107, p. 960-968.

Glen, W., 1994, On the mass extinction debates: An interview with Stephen J. Gould, *in* Glen, W., ed., The mass extinction debates: How science works in a crisis: Stanford, CA, Stanford University Press, p. 253-267.

Goldman, S.L., 1982, Modern science and western culture: The issue of time: History of European Ideas, v. 3, p. 371-401.

Gould, S.J., 1977, Ever since Darwin: Reflections in natural history: New York, W.W. Norton and Company, 285 p.

Gould, S.J., 1986, Evolution and the triumph of homology, or why history matters: American Scientist, v. 74 (January-February), p. 60-69.

Gould, S.J. 1987, Time's arrow, time's cycle: Myth and metaphor in the discovery of geological time. Cambridge, MA, Harvard University Press, 222 p.

Grieve, R.A.F., Rupert, J., Smith, J., and Therriault, A, 1995, The record of terrestrial impact cratering: GSA Today, v. 5, p. 189-196.

Hacking, I, 1983, Representing and intervening: Introductory topics in the philosophy of the natural sciences: Cambridge, Cambridge University Press, 287 p.
Hacking, I, 1999, The social construction of what?: Cambridge, MA, Harvard University Press, 261 p.
Halliday, M.A.K., 1994, An introduction to functional grammar (2nd Edition). London, Edward Arnold, 434 p.
Hildebrand, A.R., Kring, D.A., and Boynton, W.V., 1991., Chicxulub crater: A possible Cretaceous/Tertiary boundary impact crater on the Yucatan peninsula: Geology, v. 19, 867-881.
Hsu, K.J., 1994, Uniformitarianism vs. catastrophism in the extinction debates, *in* Glen, W., ed., The mass extinction debates: How science works in a crisis: Stanford, CA, Stanford University Press, p. 217-229.
Hull, D. 1973, Darwin and his Critics: The reception of Darwin's theory of evolution by the scientific community: Cambridge, MA, Harvard University Press, 473 p.
Joachims, T., 1998, Text categorization with support vector machines: Learning with many relevant features. In Machine Learning: ECML-98, Tenth European Conference on Machine Learning, pp. 137-142.
Karp, P.D., 1989, Hypothesis formation as design, *in* Shrager, J.and Langley, P. eds., Computational models of discovery and theory formation: San Francisco, Morgan Kaufmann, p. 275-315.
Kitcher, P., 1993, The advancement of science: New York, Oxford University Press, New York, 421 p.
Klein, B., and Aller, 1998, Writing across the curriculum in college chemistry: A practical bibliography: Language and learning across the disciplines, v. 2, p. 25-35.
Kuhn, T.S., 1962, The structure of scientific revolutions (1st edition): Chicago, The University of Chicago Press, 172 p.
Laudan, R., 1977, Ideas and organizations in British geology: A case study in institutional history, Isis, v. 68, p. 527-38.
Laudan, R., 1987, From mineralogy to geology: The foundations of a science, 1650-1830. Chicago, The University of Chicago Press 278 p.
LeGrand, H., 1988, Drifting continents and shifting theories: Cambridge, Cambridge University Press, 313 p.
Locke, D., 1992. Science as writing. New Haven, Yale University Press, 237 p.
Lyell, C., 1830-1833, Principles of geology (first edition): London, Murray.
Matthiessen, C., 1995, Lexicogrammatical cartography: English systems. Tokyo, Taipei & Dallas International Language Sciences Publishers, Tokyo, Taipei and Dallas, 975 p.
Mayr, E, 1985, How biology differs from the physical sciences, *in* Depew, D.J. and Weber, B.H., eds., Evolution at the crossroads: The new biology and the new philosophy of science. Cambridge, MA, MIT Press, p. 43-46.
Mayer, V.J., Champlin, R.A., Christman, R.A., and Krockover, G.H., 1980, Activity sourcebook for earth science. Columbus, Ohio, ERIC/SMEAC.
Mayer, V., and Kumano, Y., 2002, The philosophy of science and global science literacy *in* Mayer, V., ed., Global science literacy: Dordrecht, Netherlands, Kluwer Academic Publishers, p. 37-49.
Mitchell, T., 1997, Machine learning: New York, McGraw Hill, 414 p.
Moore, R., 1993, Does writing about science improve learning about science?: Journal of College Science Teaching, v, 12, p. 212-217.
National Research Council, 1996, The National Science Education Standards: Washington, D.C., National Academy, 262 p.

Ochs, E., Jacoby, S., and Gonzales, P., 1994, Interpretive journeys: How physicists talk and travel through graphic space: Configurations, v. 1, p. 151-171.
Ochs, E., and Jacoby, S., 1997, Down to the wire: The cultural clock of physicists and the discourse of consensus: Language in Society, v. 26, p. 479-506.
Oosterman, M.A., and Schmidt, M.T, eds., 1990, Earth sciences investigations: Alexandria, VA, American Geological Institute, 231 p.
Orion, N., King, C., Krockover, G.H., and Adams, P.E., 1999, The development of the earth sciences and the status of earth science education: A comparison of three case studies: Israel, England and Wales, and the United States, Part II, ICASE, v. 10, p. 13-23.
Orton, G, et. al. 1995, Collision of comet Shoemaker-Levy 9 with Jupiter observed by the NASA infrared telescope facility: Science, v. 267, p. 1277-1282.
Pardee, J.T. 1940, Ripple marks in glacial Lake Missoula: Geological Society of America Bulletin, v. 51, p. 2028-2029.
Pardee, J.T. 1942, Unusual currents in glacial Lake Missoula, Montana: Geological Society of America Bulletin, v. 53, p. 1569-1599.
Playfair, J. 1802. Illustrations of the Huttonian theory of the earth: London, Cadell and Davis, 528 p.
Popper, K.R., 1959, The logic of scientific discovery: London, Hutchinson, 479 p.
Raab, T., and Frodeman, R, 2002, What is it like to be a geologist? A phenomenology of geology and its epistemological implications: Philosophy and Geography, v. 5, p. 69-81.
Roberts, M.J., Russell, A.J., Tweed, F.S., Knudsen, O., 2001. Controls on englacial sediment deposition during the November 1996 Jokulhlaup, Skeidararjokull, Iceland: Earth Surface Processes and Landforms, v. 26, p. 935-952.
Rudolph, J.L., and Stewart, J., 1998, Evolution and the nature of science: On the historical discord and its implication for education: Journal of Research in Science Teaching, v. 35, p. 1069-1089.
Rudwick, M.J.S., 1970, The strategy of Lyell's Principles of Geology: Isis, v. 61, p. 4-33.
Rudwick, M.J.S., 1985, The meaning of fossils (2nd edition): London, MacDonald, 287 p.
Rudwick, M.J.S., 1998, Lyell and the Principles of Geology *in* Blundell, D.J., and A.C., Scott, eds., Lyell: The Past is the Key to the Present: Geological Society of London Special Publications 143, p. 3-15.
Schumm, S.A., 1991, To interpret the world: Ten ways to be wrong: Cambridge, Cambridge University Press, 133.
Schmitt, D.N., Madsen, D.B., and Lupo, K.D., 2002, Small-Mammal data on early and Middle Holocene climates and biotic communities in the Bonneville Basin, USA: Quaternary Research, v. 58, 255–260.
Sober, E, 1993, Philosophy of biology: Boulder, Co, Westview Press, 231 p.
Spray, J.G., Kelley, S.P., Rowley, D.B., 1998, Evidence for a Late Triassic multiple impact event on earth: Nature, v. 392, p. 171-173.
Toulmin, S., and Goodfield, J., 1965, The discovery of time: New York, Harper and Row, 280 p.
Turner, C., 2000, Messages in stone: Field geology in the American West, *in* Frodeman, R., ed., Earth matters: The earth sciences, philosophy and the claims of the community: Upper Saddle River, NJ, Prentice Hall, p. 51-62.
Wait, R.B., 1985, Vase for periodic colossal jokulhlaups from Pleistocene glacial lake Missoula: Geological Society of America Bulletin, v. 96, 1271-1286.
Whewell, W., 1837, History of the inductive sciences: London, John W. Parker.

Appendix: Detailed analysis of selected text excerpts

To illustrate more precisely how SFL features are realized in real texts, and how they indicate relevant aspects of the text that relate back to methodological variation, we consider here two sample passages, shown in Figures 2 and 3. The passages chosen are the conclusion sections of a physical chemistry (PhysChem) and a paleontology (Paleo) journal article, respectively. Each text has been analyzed (by hand) by breaking it into its component clauses, and marking the following components, following Halliday (1994):

- Topical themes are underlined; the theme expresses the topic of a clause, i.e., what it is 'about". Identity or similarity of theme is a strong mechanism for achieving coherence, while analysis of thematic progressions can show how a text "moves" along a storyline or through a set of related ideas.
- Conjunctions, giving the logical/causative relations between assertions, are in **boldface**,
- Nesting relations between clauses are depicted by level of indentation.
- Modal assessments, which directly quantify levels of likelihood, usuality, or necessity, are in italics, and
- Projective verb groups, whose syntactic objects are nested assertions (clauses) and which often express modality in a sort of 'grammatical metaphor', are in bold italics.

Theme

First, consider thematic development in the two passages. In the PhysChem article (Fig. 2) the topic of the passage is set in the first clause, which talks about "adsorption" of several kinds of "methylamine". This theme then splits into two "sub-stories", the first about "methylamine and dimethylamine" (which are similar), and the second about "trimethylamine". Except for the last clause in the section (which seems to be a justification of the methodology of the study), all topical themes refer either to the several methylamines studied, reactions that they participated in, or products of those reactions. This gives a strong sense of cohesion, especially in that both story-lines roughly follow parallel scripts of "chemical—reaction—result" (although the trimethylamine story contains an explicit comparison to the previous one, which also enhances cohesion).

Insert Figure 2 about here

Thematic development in the Paleo excerpt (Fig. 3), by contrast, shows a multi-focal organization. The section begins talking about the "faunal record", then focuses in on "types…of…mammals" with a shift to "the extinction of…mammals", then contrasting "abundances of … mammals" in the initially-focal "Camels Back Cave" with "Homestead Knoll". There is then a geographic shift to focus on Homestead Knoll and its geographical and climatic situation. Continuity is maintained throughout while panning over this landscape by keeping transitions small, either in semantic terms ("fauna" and "mammals" in the first four themes) or geographically (shifting between the two main locales). At the end of the passage, these multiple foci are brought together in the final two sentences, which suggest generalizations that apply to "faunas from both caves" and referring to these faunas as a single entity, "these biotic communities". Cohesion is thus achieved in this text by combining continuity (logical and geographic) of focal shift, with a concluding 'tying together' of the multiple strands of the argument.

Insert Figure 3 about here

Conjunction and Nesting

In the PhysChem passage, the sense of a single (though complex) storyline is strengthened by the kind of conjunctions used, which are all enhancing, except for the nested "also" and a single "and" near the end. Enhancing conjunctions link clauses together as part of a single information unit, by means of a logical ("although", "however") or causal ("because") link, and thus create a logically and causally coherent storyline linking the assertions. Similarly, the use of a small number of focal points for the story is also seen by the comparatively deep nesting of clauses throughout the passage.

The Paleo passage uses very different conjunctions; its multi-focal nature (noted previously) is also seen in the fact that nine out of eleven conjunctions are extensive—eight "ands" and one "moreover". The contrast with the PhysChem passage is most strikingly seen by comparing clause nesting (shown by indentation) in the two passages—the Paleo passage is nearly flat, while the PhysChem passage has quite a complex hierarchy of nested clauses. We suggest that the flatness of the Paleo passage gives it a sort of "panoramic" text style, with an effect of looking around at a geographic/conceptual landscape, rather than following a single causally determined storyline as in the PhysChem passage.

Modality

We now consider how the two passages relate to the likelihood and typicality of the phenomena and results they describe. Noteworthy in the PhysChem passage is the almost exclusive use of apparently stereotypical projective clauses to hedge on certainty ("appear", "seems"). The uniformity of phrasing argues that such usages constitute predefined 'templates' in the genre, and do not substantially contribute to the argument being made in the passage.

Modal assessment in the Paleo passage differs, however, in the way it is expressed. About half of the modal assessments in this passage are 'implicit', realized by modal adjuncts, such as "probably", "always", and "likely", rather than by projective verbs such as "appear"; the greater variety of modal structures argues that they are more semantically significant in constructing the argument. For example, the double modal "probably always" establishes universality for the "lake-effect", softened for plausibility by the "probably". This is key to drawing the conclusion that "this lake-edge contrast has always been wetter…" This illustrates that clear assessments of likelihood and typicality are needed for historically based scientific argumentation.

TABLE 1. TYPES OF COMMUNICATION BETWEEN SCIENTISTS AND THEIR AUDIENCES

Types of Communication		Audience		
		Collaborators	Colleagues	Students
Speech		Lab Meetings, Field Research	Conference Talks	Lectures
Writing	Informal	Research notes	Correspondence	Lecture notes
	Formal	Lab reports	Published articles	Textbooks

TABLE 2. JOURNALS INCLUDED IN THE CORPUS-BASED STUDY

Science Type*	Field	Journal	# articles	Avg. words/article
Experimental	Physics	*Phys. Letters A*	132	2339
		Physical Rev. Letters	114	2545
	Organic Chemistry	*Heterocycles*	231	3580
		Tetrahedron	151	5057
	Physical Chemistry	*J. Phys. Chem. A*	121	4865
		J. Phys. Chem. B	71	5269
Historical	Geology	*J. Geology*	93	4891
		J. Metamorphic Geol.	108	5024
	Evolutionary Biology	*Biol. J. Linnean Soc.*	191	4895
		Human Evolution	169	4223
	Paleontology	*Paleontologica Elec.*	111	4132
		Quaternary Research	113	2939

Note: All articles in this corpus are from 2003.
*The authors classified the journals into their science type (historical or experimental) prior to the computer-aided investigation.

Figure 1. A plot of classification effectiveness as measured by cross validation accuracy (Acc.). The open circles represent mean classification accuracies in a specific category (Same, Historical, Experimental, and Different) and the small squares represent measurements for each pair of journals compared in a specific category.

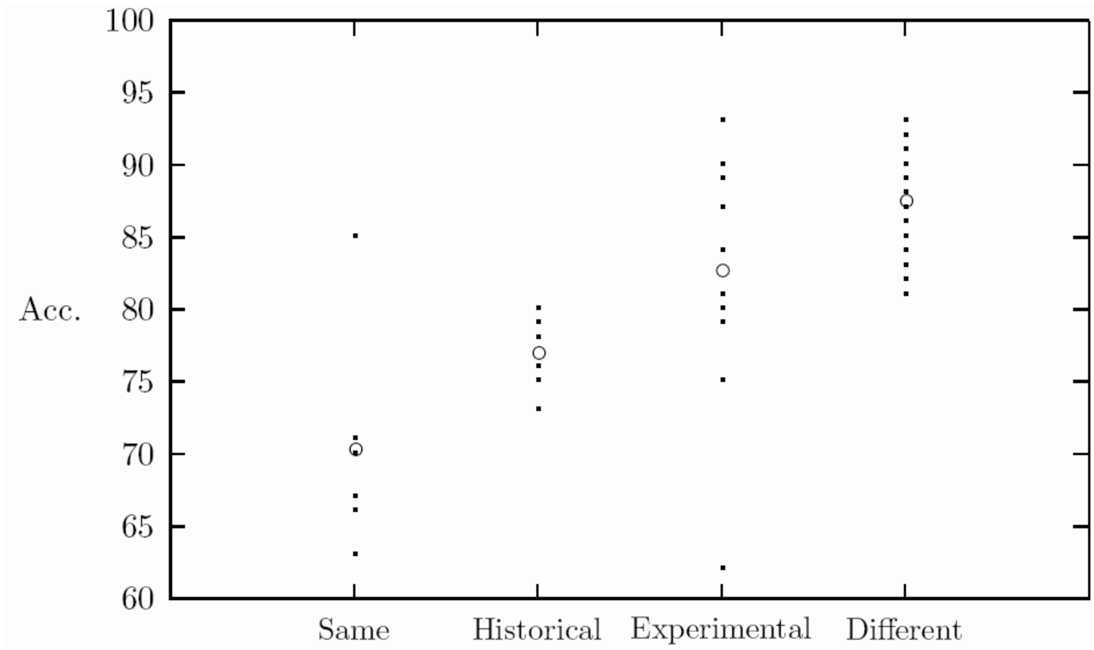

Figure 2. The conclusion section from April J. Carman, Linhu Zhang, Jason L. Liswood, and Sean M. Casey, "Methylamine Adsorption on and Desorption from Si(100)", J. Phys. Chem. B 2003, 107, 5491-5502. The text is split into individual clauses, with topical themes underlined, conjunctions in boldface, modal finites and adjuncts in italics, and projective verb groups in bold italics. Indentation indicates hierarchical organization of clauses and subclauses.

- The adsorption of several simple methylamines on the Si-(100)-(2x1) surface has been investigated using Auger electron .spectroscopy (AES), thermal desorption spectroscopy, and lowenergy electron diffraction
- Both methylamine and dimethylamine ***appear***
  - to undergo mostly dissociative adsorption on this surface at room temperature,
- **although**
  - trapping into a molecular adsorption well ***appears***
    - to occur to a limited extent for both molecules, as observed by detection of a parent desorption channel in the 410-430-K temperature range.
- The major desorbing thermal decomposition product for both of these species ***seems***
  - to come from a β-hydrogen elimination-like reaction of a surface-bound $NH_x(CH_3)_{2-x}$ (x = 1 or 0) to form an imine, a reaction
    - that is ***likely*** to be analogous to the formation of ethylene from surface-bound ethyl groups.
- This **also** results in formation of molecular hydrogen, the other main desorbing product observed for both of these species.
- Trimethylamine, **however**, ***appears***
  - to undergo mostly molecular adsorption on this surface.
- By comparison to the AES results from the adsorption of methyl iodide on Si(100), it ***was concluded that***
  - the initial surface saturation coverage of trimethylamine on Si(100) is 0.26 monolayers,
  - **while**
  - both methylamine and dimethylamine ***appear***
    - to saturate at about 0.48 monolayers.
- The difference in saturation coverage in the case of trimethylamine ***appears to be***
  - **because**
    - The molecule is stuck in the physisorption well.
- This results in blocking of adjacent bonding sites, *most likely* by delocalized surface electronic accommodation of the trimethylamine-silicon dative bond
- **and**
- leads to a coverage approximately half of that observed for methylamine and dimethylamine.
- Ab initio and density functional calculations, even at relatively low levels, ***seem***
  - to capture the energetics of these surface reactions within about 15% (about 15-25 kJ/mol) compared to kinetic modeling (using computed and standard frequency factors) of thermal desorption results.

Figure 3. The conclusion section from Dave N. Schmitt, David B. Madsen, and Karen D. Lupo, “Small-Mammal Data on Early and Middle Holocene Climates and Biotic Communities in the Bonneville Basin, USA”, Quaternary Research 58, 255–260 (2002). The text is marked up as in Figure 2.

The early faunal record at Camels Back Cave provides data on early Holocene environments and mammalian responses to middle Holocene desertification in the Bonneville Basin.	
This record adds to the growing body of data on small mammal histories in the Great Basin	
and	
may prove useful in modern wildlife management issues (e.g., Lyman, 1996), especially those concerning the potential impacts of future climatic change on mammal populations in the arid west.	
The types and variety of early Holocene mammals ***suggest that***	
	the Camels Back Ridge vicinity was cool and moist
	and
	supported grasses and stands of *Artemisia* (likely *A. tridentata*),
and	
the extinction of small mammals adapted to mesic contexts just prior to 8000 14C yr B.P. attests to the extreme aridity of the middle Holocene.	
The relative abundances of these early Holocene mammals are less than those at Homestead Cave	
and	
It ***appears that***	
	areas surrounding Camels Back Cave were not quite as cool.
This is not surprising **however**,	
Since	
	Homestead Knoll is 120 km to the north
	and
	areas immediately south of the cave encompass a more upland setting than southern Camels Back Ridge.
Moreover,	
Homestead Cave is adjacent the Great Salt Lake (e.g., Madsen *et al.*, 2001, Figure 3)	
and	
	as
	Pacific storm systems make their way across the area,
the “lake-effect” enhancement of these storms has *probably always* come into play,	
and	
It is *likely* that	
	this lake-edge context has *always* been wetter and slightly cooler than the southern Great Salt Lake Desert.
Regardless of these subtle differences, faunas from both caves strongly ***suggest that***	
	middle Holocene desertification brought forth significant changes in regional plants and animals, including a rather rapid transition to xerophytic shrub communities dominated by *Sarcobatus vermiculatus* and *Atriplex* sp. and an overall decline in mammalian taxonomic richness.
These biotic communities have dominated the Camels Back Cave vicinity for the last ca. 8000 14C years	
and	
probably were similar to the native desert habitats surrounding Camels Back Ridge today.	